



**HAL**  
open science

# Using Covert Response Activation to Test Latent Assumptions of Formal Decision-Making Models in Humans

Mathieu Servant, Corey White, Anna Montagnini, Boris Burle

► **To cite this version:**

Mathieu Servant, Corey White, Anna Montagnini, Boris Burle. Using Covert Response Activation to Test Latent Assumptions of Formal Decision-Making Models in Humans. *Journal of Neuroscience*, 2015, 35 (28), pp.10371-10385. 10.1523/JNEUROSCI.0078-15.2015 . hal-01310094

**HAL Id: hal-01310094**

**<https://amu.hal.science/hal-01310094>**

Submitted on 1 May 2016

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Using Covert Response Activation to Test Latent Assumptions of Formal Decision-Making Models in Humans

Mathieu Servant,<sup>1,2</sup> Corey White,<sup>3</sup> Anna Montagnini,<sup>2</sup> and Boris Burle<sup>1</sup>

<sup>1</sup>Laboratoire de Neurosciences Cognitives, Fédération de Recherche 3C, Case C, Aix-Marseille Université, Centre National de la Recherche Scientifique, 13331 Marseille cedex 3, France, <sup>2</sup>Institut de Neurosciences de la Timone, Aix-Marseille Université, Centre National de la Recherche Scientifique, Faculté de Médecine, 13005 Marseille, France, and <sup>3</sup>Brain and Behavior Laboratory, Department of Psychology, Syracuse University, Syracuse, New York 13244

Most decisions that we make build upon multiple streams of sensory evidence and control mechanisms are needed to filter out irrelevant information. Sequential sampling models of perceptual decision making have recently been enriched by attentional mechanisms that weight sensory evidence in a dynamic and goal-directed way. However, the framework retains the longstanding hypothesis that motor activity is engaged only once a decision threshold is reached. To probe latent assumptions of these models, neurophysiological indices are needed. Therefore, we collected behavioral and EMG data in the flanker task, a standard paradigm to investigate decisions about relevance. Although the models captured response time distributions and accuracy data, EMG analyses of response agonist muscles challenged the assumption of independence between decision and motor processes. Those analyses revealed covert incorrect EMG activity (“partial error”) in a fraction of trials in which the correct response was finally given, providing intermediate states of evidence accumulation and response activation at the single-trial level. We extended the models by allowing motor activity to occur before a commitment to a choice and demonstrated that the proposed framework captured the rate, latency, and EMG surface of partial errors, along with the speed of the correction process. In return, EMG data provided strong constraints to discriminate between competing models that made similar behavioral predictions. Our study opens new theoretical and methodological avenues for understanding the links among decision making, cognitive control, and motor execution in humans.

**Key words:** accumulation models; decision making; electromyography; electrophysiology; partial errors

## Significance Statement

Sequential sampling models of perceptual decision making assume that sensory information is accumulated until a criterion quantity of evidence is obtained, from where the decision terminates in a choice and motor activity is engaged. The very existence of covert incorrect EMG activity (“partial error”) during the evidence accumulation process challenges this longstanding assumption. In the present work, we use partial errors to better constrain sequential sampling models at the single-trial level.

## Introduction

Model-based cognitive neuroscience aims to use formal models of cognition to inform neurophysiological data while reciprocally using neurophysiological data to better constrain formal models.

Received Jan. 8, 2015; revised May 19, 2015; accepted June 9, 2015.

Author contributions: M.S., A.M., and B.B. designed research; M.S. performed research; M.S., C.W., A.M., and B.B. contributed unpublished reagents/analytic tools; M.S., C.W., A.M., and B.B. analyzed data; M.S., C.W., A.M., and B.B. wrote the paper.

This work was supported by the European Research Council under the European Community's Seventh Framework Program (FP/2007–2013 Grant Agreement 241077) and by the Centre National de la Recherche Scientifique. We thank Bram Zandbelt and an anonymous reviewer for helpful comments.

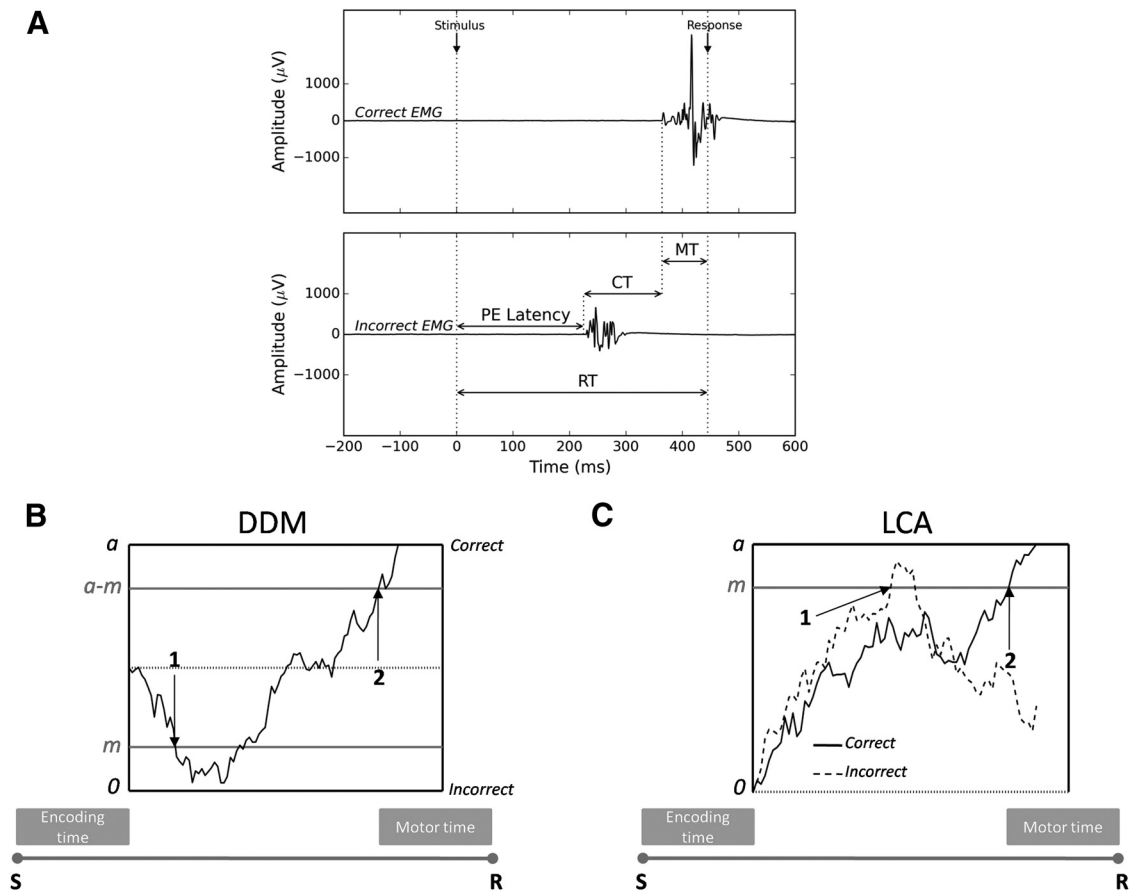
The authors declare no competing financial interests.

Correspondence should be addressed to Boris Burle, Laboratoire de Neurosciences Cognitives, Fédération de Recherche 3C, case C, Aix-Marseille Université, CNRS, 3 Place Victor Hugo, 13331 Marseille cedex 3, France. E-mail: boris.burle@univ-amu.fr.

DOI:10.1523/JNEUROSCI.0078-15.2015

Copyright © 2015 the authors 0270-6474/15/3510371-15\$15.00/0

This multidisciplinary approach has led to substantial advances for understanding higher cognitive functions (Forstmann et al., 2011). In particular, perceptual decision making has benefited from interactive relationships between sequential sampling models and single-unit electrophysiology (Purcell et al., 2010; Heitz and Schall, 2012), fMRI (van Maanen et al., 2011; White et al., 2012), and EEG (Ratcliff et al., 2009; O'Connell et al., 2012). Although fMRI and EEG allow researchers to study decision making in humans at a systems level, both techniques present important limitations. The low temporal resolution of fMRI prevents any characterization of decision signal dynamics and most EEG studies rely on averaging procedures known to generate temporal distortions (Meyer et al., 1988; Burle et al., 2008). Although single-trial analysis techniques are currently being developed (Sadja et al., 2011), this approach is limited by the very small signal-to-noise ratio of the EEG coupled with the high variety of



**Figure 1.** EMG partial errors and the sequential sampling model framework. **A**, EMG activity (in  $\mu V$ ) of the muscles involved (Correct EMG, top) and not involved (Incorrect EMG, bottom) in the required response as a function of time (in ms) from stimulus onset. Partial motor activity in the incorrect EMG channel precedes the correct response. RT, Reaction time (from stimulus onset to the mechanical response); PE, partial error; CT, correction time (from the incorrect EMG activation to the correct one); MT, motor time (from the correct EMG activation to the mechanical response). **B**, Extended DDM. EMG bounds were incorporated within the response selection accumulator at locations  $m$  (incorrect EMG activation) and  $a-m$  (correct EMG activation). EMG bounds do not correspond to an actual choice. Evidence continues to accumulate until standard decision termination bounds 0 and  $a$  are reached. Therefore, part of the MT overlaps with the decision time. The decision sample path represents a partial error trial. Arrows correspond to EMG events: 1 = onset of the partial error, 2 = onset of the correct EMG burst. *S*, Stimulus; *R*, mechanical response. **C**, Extended LCA model. The two competing response accumulators are superimposed for convenience. An EMG bound was incorporated within each response accumulator at location  $m$ . The two decision sample paths represent a partial error trial. The sample path favoring the incorrect response (dashed sample path) hits the EMG bound  $m$ , but the sample path favoring the correct response (solid sample path) finally reaches the decision termination bound  $a$  and wins the competition. Arrows correspond to EMG events: 1 = onset of the partial error, 2 = onset of the correct EMG burst.

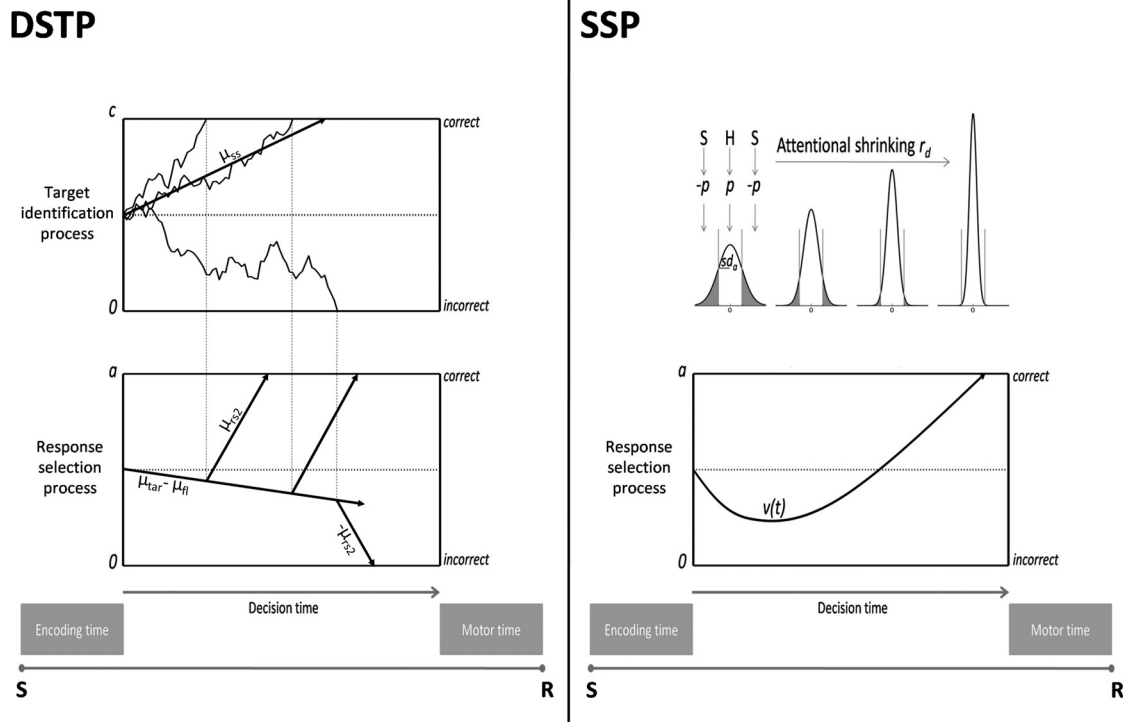
noise sources. EMG provides an interesting alternative. Albeit restricted to the peripheral level, EMG has an excellent signal-to-noise ratio and can be analyzed on a trial-to-trial basis (Coles et al., 1985). In between-hand choice reaction time (RT) tasks, previous work has shown covert incorrect EMG activation in a portion of trials in which the correct response was finally given (Fig. 1; Coles et al., 1985; Burle et al., 2014). This phenomenon, termed “partial error,” challenges sequential sampling models of perceptual decision making that assume that motor activity is engaged only once a decision threshold is reached (Ratcliff and Smith, 2004; Bogacz et al., 2006). Although the models have recently been extended to account for more ecological choice situations requiring filtering out of irrelevant sensory information (Liu et al., 2008; Hübner et al., 2010; White et al., 2012; White et al., 2011), they preserve the longstanding assumption of independence between decision and motor processes. The very existence of EMG partial errors invites a reappraisal of this assumption and requires a theory explaining how, why, and when discrete motor events sometimes emerge in the EMG during the decision process. The purpose of the present study is to use EMG data to better constrain sequential sampling models while reciprocally using the models to inform EMG data. To this aim, we collected behav-

ioral and EMG data from a flanker task, a standard paradigm to investigate decisions about relevance. After fitting different classes of models to each behavioral dataset, comparison between EMG signals and simulated decision sample paths revealed dependence between decision formation and response execution because part of the EMG activity was determined by the evolving decision variable. Incorporating bounds for EMG activation into the sequential sampling framework allowed the models to capture a range of EMG effects. In return, EMG data provided strong constraints to discriminate between competing models indiscernible on behavioral grounds.

## Materials and Methods

### Sequential sampling models

All classes of sequential sampling models share the assumption that sensory evidence is accumulated to some criterion level or bound, from where the decision terminates in a choice and the response is executed. In this section, we introduce two major classes of models that have proven to account for behavioral and neural data of perceptual decision tasks: the drift diffusion model (DDM) (Ratcliff, 1978; Gold and Shadlen, 2007; Ratcliff et al., 2007; Ratcliff and McKoon, 2008) and the leaky competing accumulator model (LCA) (Usher and McClelland, 2001; Purcell et al., 2012; Schurger et al., 2012).



**Figure 2.** Diffusion models incorporating selective attention mechanisms. An application to an incongruent flanker condition is provided as a working example. Left, DSTP model. Dashed lines joining the two accumulators highlight the effect of three target identification sample paths on response selection. Right, SSP model. S, Stimulus; R, mechanical response. See text for details.

*DDM.* The DDM assumes continuous and perfect accumulation of the difference between noisy samples of sensory evidence supporting two response alternatives. Accumulation begins from starting point  $z$  toward one of two decision termination bounds  $a$  (correct choice) and  $0$  (incorrect choice) according to the following:

$$dx = vdt + sdW$$

Where  $dx$  represents the change in accumulated evidence  $x$  for a small time interval  $dt$ ,  $v$  is the drift rate (i.e., the average increase of evidence in favor of the correct choice) and  $sdW$  denotes Gaussianly distributed white noise with mean 0 and variance  $s^2 dt$ . The magnitude of noise  $s$  acts as a scaling parameter and is usually fixed at 0.1 (Ratcliff and Smith, 2004; Ratcliff and McKoon, 2008). The RT is the time required to reach a decision termination bound plus residual processing latencies such as sensory encoding and motor execution (residual components are generally combined into one parameter with mean time  $Ter$ ).

In the standard DDM, the drift rate is constant over time (Ratcliff and McKoon, 2008). However, the model has recently been enriched by attentional control mechanisms that weight sensory evidence in a flexible and goal-directed way (Hübner et al., 2010; White et al., 2011), allowing capture of behavioral performance in more ecological choice situations requiring filtering of irrelevant information. Such situations have traditionally been studied with the flanker paradigm (Eriksen and Eriksen, 1974). Subjects are instructed to press a right or a left button according to a central symbolic target (e.g., a left response to the letter S) and a right response to the letter H). The target is flanked by distractors calling for the same (HHHHH or SSSSS) or the opposite response (SSHSS or HSHHH). Incongruent flankers produce interference, leading to slower and less accurate responses. Moreover, distributional analyses have revealed that interference is larger early in the course of processing because errors are concentrated in the leftmost part of the RT distribution (Gratton et al., 1988). Extensions of the DDM explain performance in this task by an increase of attentional selectivity during stimulus processing, resulting in time-varying decision evidence. Two model extensions have been proposed, differing in whether selective attention proceeds in a discrete or continuous manner (Fig. 2).

The dual-stage two-phase model of selective attention (DSTP) (Hübner et al., 2010) integrates early and late selection theories of attention (Broadbent, 1958; Deutsch and Deutsch, 1963) into the DDM framework. An early stage of low selectivity drives a response selection diffusion process in phase 1. In the context of a flanker task, the drift rate for this first phase is the sum of the attention-weighted sensory evidence from the target  $\mu_{tar}$  and flanking items  $\mu_{fl}$ . If flankers are incongruent,  $\mu_{fl}$  is negative, which decreases the net drift rate and enhances the probability of making fast errors. In parallel, a second diffusion process with drift rate  $\mu_{ss}$  selects a target based on identity, which is more time consuming but highly selective. Because two diffusion variables are racing, processing can be categorized into two main schemes. First, if the response selection process hits a termination bound before the target identification process finishes, the model reduces to a standard DDM. Second, if the target identification process wins the race, from that moment on, the identified target drives response selection in phase 2 and the drift rate increases discretely from  $\mu_{tar} \pm \mu_{fl}$  to  $\pm \mu_{rs2}$ . If target identification is correct, then  $\mu_{rs2}$  is positive and countervails early incorrect activations in incongruent trials, explaining the improved accuracy of slower responses (see Fig. 2, left, for an illustration of this processing scheme). If target identification is incorrect,  $\mu_{rs2}$  is negative and the model generates a slow perceptual error. To summarize, the time-varying drift rate  $v(t)$  that drives the response selection process of the DSTP is given by a piecewise function of the following form:

$$\text{Phase 1: } v(t) = \mu_{fl} \pm \mu_{tar}; \text{ Phase 2: } v(t) = \pm \mu_{rs2}$$

Alternatively, the shrinking spotlight model (SSP) (White et al., 2011) incorporates principles of a “zoom-lens” model of visual spatial attention (Eriksen and St James, 1986) into the DDM framework. Attention is likened to a spotlight of a size that can be varied continuously depending on task demands. For example, in the flanker task, the attention spotlight is diffuse early in a given trial and progressively shrinks to focus on the target (Fig. 2, right). Because attentional resources are assumed to be fixed, the spotlight is modeled as the integral of a normalized Gaussian distribution centered on the target. Attention shrinking is performed by decreasing the SD  $sd_a$  of the Gaussian at a linear rate  $r_d$ . Each item in the

display is one unit wide with the exception of the outer flankers that receive the excess of attention, ensuring that the spotlight area always sum up to 1. Assuming a standard 5-item display with the target being centered at 0, the quantity of attention allocated to the target, right inner flanker and right outer flanker over time (the left flankers are symmetrical) is given as follows:

$$a_{target}(t) = \int_{-.5}^{.5} N[0, sd_a(t)];$$

$$a_{inner}(t) = \int_{.5}^{1.5} N[0, sd_a(t)]; a_{outer}(t) = \int_{1.5}^{\infty} N[0, sd_a(t)]$$

Where  $sd_a(t) = sd_a(0) - r_d(t)$

The target and each flanker contribute the same quantity of sensory evidence, referred to as perceptual input  $p$  (incongruent flankers have a negative perceptual input). In each time step, perceptual inputs are weighted by the allocated quantity of attention that defines the continuously evolving drift rate  $v(t)$ :

$$v(t) = p_{target}a_{target}(t) + 2p_{inner}a_{inner}(t) + 2p_{outer}a_{outer}(t)$$

For incongruent trials,  $v(t)$  initially favors the incorrect response and progressively heads toward the correct bound as the spotlight shrinks. Notice that  $v(t)$  is constant for congruent trials because attention weights always sum up to 1.

The DSTP and SSP capture RT distributions and accuracy data from flanker tasks under a wide variety of experimental manipulations (Hübner et al., 2010; White et al., 2011). This new generation of diffusion models also explains particular interactions between flanker congruency and well known psychological laws (Servant et al., 2014). However, the DSTP and SSP so largely mimic each other on behavioral grounds that it has proven difficult to determine the superiority of one model over another. EMG indices will prove to be useful for constraining their evidence accumulation dynamics and falsifying their latent processing assumptions.

**LCA model.** The LCA model assumes sensory evidence integration into separate accumulators, each one representing a population of neurons favoring one of the  $N$  possible response alternatives. It is neurally inspired in the sense that it takes into account leakage of integration and assumes mutual inhibition between accumulators (Usher and McClelland, 2001). Assuming  $N = 2$ , accumulation processes  $x_1$  and  $x_2$  for correct and incorrect response alternatives start at 0 and drift toward a decision bound  $a$  according to the following:

$$\begin{cases} dx_1 = (-kx_1 - wx_2 + I_1)dt + sdW_1 \\ dx_2 = (-kx_2 - wx_1 + I_2)dt + sdW_2 \end{cases}$$

Where  $I_i$  represents the average increase of evidence for each response alternative,  $-kx_i$  leakage of integration and  $-wx_i$  mutual inhibition. Notice that the amount of mutual inhibition exerted by an accumulator on the competing one grows linearly with its own activation. Likewise, the leakage of one accumulator grows linearly with its own activation. Because neural firing rates are never negative, a threshold-linear function  $f(x_i)$  is applied to each integration process such that  $f(x_i) = 0$  for  $x_i < 0$  and  $f(x_i) = x_i$  for  $x_i \geq 0$ . The stochastic component of the LCA is similar to that of the DDM.

The standard LCA generally predicts slower errors than correct responses. Like the standard DDM, the only way for the LCA to produce faster errors than correct responses is to add across-trial variability in the starting points of the accumulators (Usher and McClelland, 2001; Ratcliff and Smith, 2004). In the context of a flanker task, this additional assumption would lead to equivalent increases in fast errors across congruency conditions, contrary to what is usually observed (Gratton et al., 1988; Hübner et al., 2010; White et al., 2011; Servant et al., 2014). Therefore, the standard LCA cannot account for behavioral performance in the flanker task for the same reason as for the standard DDM (White et al.,

2011). We thus developed an SSP-LCA that incorporates the shrinking spotlight attentional mechanism developed by White et al. (2011b) into the LCA framework. Contrary to the SSP, the perceptual input  $p$  of any item in the display is always positive and feeds into different response accumulators depending on the congruency condition. The SSP-LCA is thus defined as follows:

$$\begin{cases} dx_1 = (-kx_1 - wx_2 + I_1(t))dt + sdW_1 \\ dx_2 = (-kx_2 - wx_1 + I_2(t))dt + sdW_2 \end{cases}$$

Where

$$\text{incongruent} \begin{cases} I_1(t) = p_{target}a_{target}(t) \\ I_2(t) = 2p_{inner}a_{inner}(t) + 2p_{outer}a_{outer}(t) \end{cases}$$

$$\text{congruent} \begin{cases} I_1(t) = p_{target}a_{target}(t) + 2p_{inner}a_{inner}(t) + 2p_{outer}a_{outer}(t) \\ I_2(t) = 0 \end{cases}$$

The DSTP could also be rewritten in the LCA framework. However, the model has seven free parameters. Incorporating leakage and mutual inhibition principles would result in at least six additional free parameters ( $k_1$  and  $w_1$  for the target identification process,  $k_2$  and  $w_2$  for the response selection process in phase 1,  $k_3$  and  $w_3$  for the response selection process in phase 2). This would strongly increase the flexibility of the model and the risk of parameter tradeoffs. For this reason, we analyzed a simpler DSTP-LCA in which leakage and mutual inhibition are constrained to be identical across processes. This model has nine free parameters and is defined as follows:

$$\begin{aligned} \text{response selection} & \begin{cases} dx_1 = (-kx_1 - wx_2 + I_1(t))dt + sdW_1 \\ dx_2 = (-kx_2 - wx_1 + I_2(t))dt + sdW_2 \end{cases} \\ \text{target identification} & \begin{cases} dx_3 = (-kx_3 - wx_4 + I_3(t))dt + sdW_3 \\ dx_4 = (-kx_4 - wx_3 + I_4(t))dt + sdW_4 \end{cases} \end{aligned}$$

where

$$I_3 = \mu_{ss} \quad I_4 = 0$$

$$\text{incongruent} \begin{cases} \text{Phase 1: } I_1(t) = \mu_{tar} & I_1(t) = \mu_{fl} \\ \text{Phase 2: } I_1(t) = \mu_{rs2} \text{ or } 0 & I_2(t) = \mu_{rs2} \text{ or } 0 \\ \text{depending of which process } x_3(t) \text{ or } x_4(t) \\ \text{has won the race} \end{cases}$$

$$\text{congruent} \begin{cases} \text{Phase 1: } I_1(t) = \mu_{tar} + \mu_{fl} & I_2(t) = 0 \\ \text{Phase 2: } I_1(t) = \mu_{rs2} \text{ or } 0 & I_2(t) = \mu_{rs2} \text{ or } 0 \\ \text{depending of which process } x_3(t) \text{ or } x_4(t) \\ \text{has won the race} \end{cases}$$

### Participants

Twelve students (mean age = 22.9 years, SD = 4.1, 4 males and 8 females) with normal or corrected-to-normal vision participated in the study and were paid 10 Euros/h. They were not aware of the purpose of the experiment. This experiment was approved by the ethical committee of the Aix-Marseille University and by the Comité de Protection des Personnes Sud Méditerranée 1 (approval #1041). Participants gave their informed written consent according to the Declaration of Helsinki.

### Stimuli and apparatus

Each stimulus array consisted in five white letters (S or H) presented on the horizontal midline of a  $12.18^\circ \times 9.15^\circ$  black field. The target letter was always presented at the center of the screen. A congruent stimulus array consisted of the target letter flanked by identical letters (SSSSS or HHHHH). An incongruent array consisted of the target letter flanked by the alternative letters (SSHSS or HSHSH). Each letter subtended  $0.7^\circ$  of visual angle with  $0.6^\circ$  separation between the letters. Responses were made by pressing either a left or a right button with the corresponding thumb. Buttons were fixed on the top of two plastic cylinders (3 cm in diameter, 7 cm in height) separated by a distance of 10 cm. Button clo-

tures were transmitted to the parallel port of the recording PC to reach high temporal precision.

*Procedure*

Participants were tested in a dark and sound-shielded Faraday cage and seated in a comfortable chair 100 cm in front of a CRT monitor with a refresh rate of 75 Hz. They were instructed to respond quickly and accurately to a central letter (H or S) and ignore flanking letters. Half of the participants gave a left-hand response to the letter H and a right-hand response to the letter S. This mapping was reversed for the other half. Stimulus presentation and collection of data were controlled using components of Psychopy (Peirce, 2007). Participants first performed 96 practice trials and worked through 15 blocks of 96 trials. Practice trials were excluded from analyses. All types of trials (SSSSS, HHHHH, SSHSS, HSHHH) were equiprobable and presented in a pseudorandom order so that first-order congruency sequences (i.e., congruent-incongruent CI, CC, IC, II) occurred the same number of times. On every trial, a stimulus array was presented and remained on the screen until a response was given. The next trial started 1000 ms after stimulus offset.

*Model fitting*

The DSTP, SSP, DSTP-LCA, and SSP-LCA were fit to each individual dataset using standard quantile-based methods. Basically, the models were simulated using a random walk numerical approximation and a 0.001 s time step to produce probabilities and RT quantiles (0.1, 0.3, 0.5, 0.7, 0.9) for correct and error responses. Those values were compared against data through a  $G^2$  likelihood ratio statistic (Ratcliff and Smith, 2004) as follows:

$$G^2 = 2 \left( \sum_{i=1}^2 n_i \sum_{j=1}^X p_{ij} \log \frac{p_{ij}}{\pi_{ij}} \right)$$

The outer summation over  $i$  extends over the two flanker congruency conditions.  $n_i$  is the number of valid trials per condition. The variable  $X$  represents the number of bins bounded by RT quantiles for each distribution pair of correct and error responses. Because errors are scarce in the congruent condition, we only took into account their median RT. Therefore,  $X = 8$  (6 bins for correct responses and 2 bins for errors) for the congruent condition and  $X = 12$  otherwise.  $p_{ij}$  and  $\pi_{ij}$  are, respectively, the observed and predicted proportions of responses in bin  $j$  of condition  $i$ . No parameter was allowed to vary between congruency conditions apart from the sign of  $p$  for the SSP and the sign of  $\mu_n$  for the DSTP. The magnitude of noise  $s$  served as a scaling parameter and was set to 0.1 in all models (Ratcliff and Smith, 2004). We assumed unbiased starting points of diffusion processes given that all types of trials were equiprobable. The  $G^2$  statistic was minimized by a SIMPLEX routine (Nelder and Mead, 1965) to obtain best-fitting model parameters. Eighty thousand trials were simulated for each congruency condition and minimization cycle (the following model simulations always contain 80,000 trials per congruency condition). Each model was fit several times using the best-fitting values from the previous fits as the new starting values for the next run. Different sets of initial parameter values were used to ensure that the SIMPLEX gradient descent does not reach a local minimum. For the DSTP and SSP diffusion models, initial values for each parameter were randomly chosen from uniform distributions bounded by the minimum and maximum of previous reported fits of the models (Hübner et al., 2010; White et al., 2011; Hübner and Töbel, 2012; Servant et al., 2014). For the DSTP-LCA and SSP-LCA models, we broadened these bounds to take into account previous reported fits of the standard LCA. Because the DSTP-LCA and SSP-LCA have never been fit to data, we checked our results by running a global optimizer (differential evolution; Storn and Price, 1997). Both methods gave approximately similar results.

The  $G^2$  likelihood ratio statistic was considered as a relative measure of fit quality and was completed by a BIC statistic (Schwarz, 1978) that penalizes models as a function of their number of free parameters  $f$  (the best model is the one with the smaller BIC) as follows:

$$BIC = G^2 + f \log \sum_{i=1}^2 n_i$$

Where  $n_i$  is the number of valid trials per condition. The goodness-of-fit of the models can also be appreciated in Figure 3, where data and model predictions averaged over participants are represented as quantile probability functions (QPFs). QPFs are constructed by plotting RT quantiles ( $y$ -axis) of the distributions of correct and incorrect responses for each experimental condition against the corresponding response type proportion ( $x$ -axis).

*EMG recordings and signal processing*

The EMG activities of the flexor pollicis brevis of both hands were recorded by means of 2 pairs of Ag/Cl electrodes (BIOSEMI Active-Two electrodes) glued 2 cm apart on the thenar eminence (sampling rate = 2048 Hz). This activity was filtered after acquisition (high-pass = 10 Hz). The EMG signal was processed with BrainAnalyser (Brain Products) and custom programs written in Python (www.python.org). The onset of EMG activity was detected by running the EMG Onset Search method as implemented in BrainAnalyser (Brain Products). The method is based on an algorithm developed by Van Boxtel et al. (1993). For each trial, the mean and SD amplitude of a 200 ms prestimulus baseline were first computed. EMG onset was then detected when the amplitude of the signal exceeds  $n \times SD$ ,  $n$  being manually adapted per participant as a function of signal-to-noise. A backward search for a sign change in the second derivative peak was finally performed to refine the precision of the detection. Because EMG onset detection algorithms are imperfect (Van Boxtel et al., 1993), inaccurate EMG onset markers were manually corrected after visual inspection. Indeed, it has been shown that visual inspection remains the most accurate technique (Stauder et al., 2001), especially for detecting small changes in EMG activity such as partial errors. The experimenter who corrected the markers was unaware of the type of mapping the EMG traces corresponded to. Trials were classified “correct” or “error” depending on whether button closure took place in the correct or incorrect EMG channel. Among correct trials, trials in which an incorrect EMG activity preceded the correct one (“partial error trials”; Fig. 1A) were dissociated from “pure correct trials.” “Other” trials (10.7%) containing tonic muscular activity or patterns of EMG bursts not categorizable in the above classes were discarded from analyses.

*Data analyses*

Unless specified, all statistical analyses were performed by means of repeated-measures ANOVAs with flanker congruency as factor. Proportions were arc-sine transformed before being submitted to ANOVA to stabilize their variance (Winer, 1971). EMG analyses focused on partial errors and associated chronometric indices, the latency of partial errors, and their correction time (time from the incorrect EMG activation to the correct one; Fig. 1A). Partial error EMG surfaces were measured as follows. For each subject and congruency condition, EMG bursts corresponding to partial errors were rectified, baseline corrected, averaged, and time locked to their onset. The cumulative sum of the averaged partial error burst was then computed and the 90<sup>th</sup> percentile was taken as an estimate of the surface. The same procedure was applied to EMG bursts of overt responses.

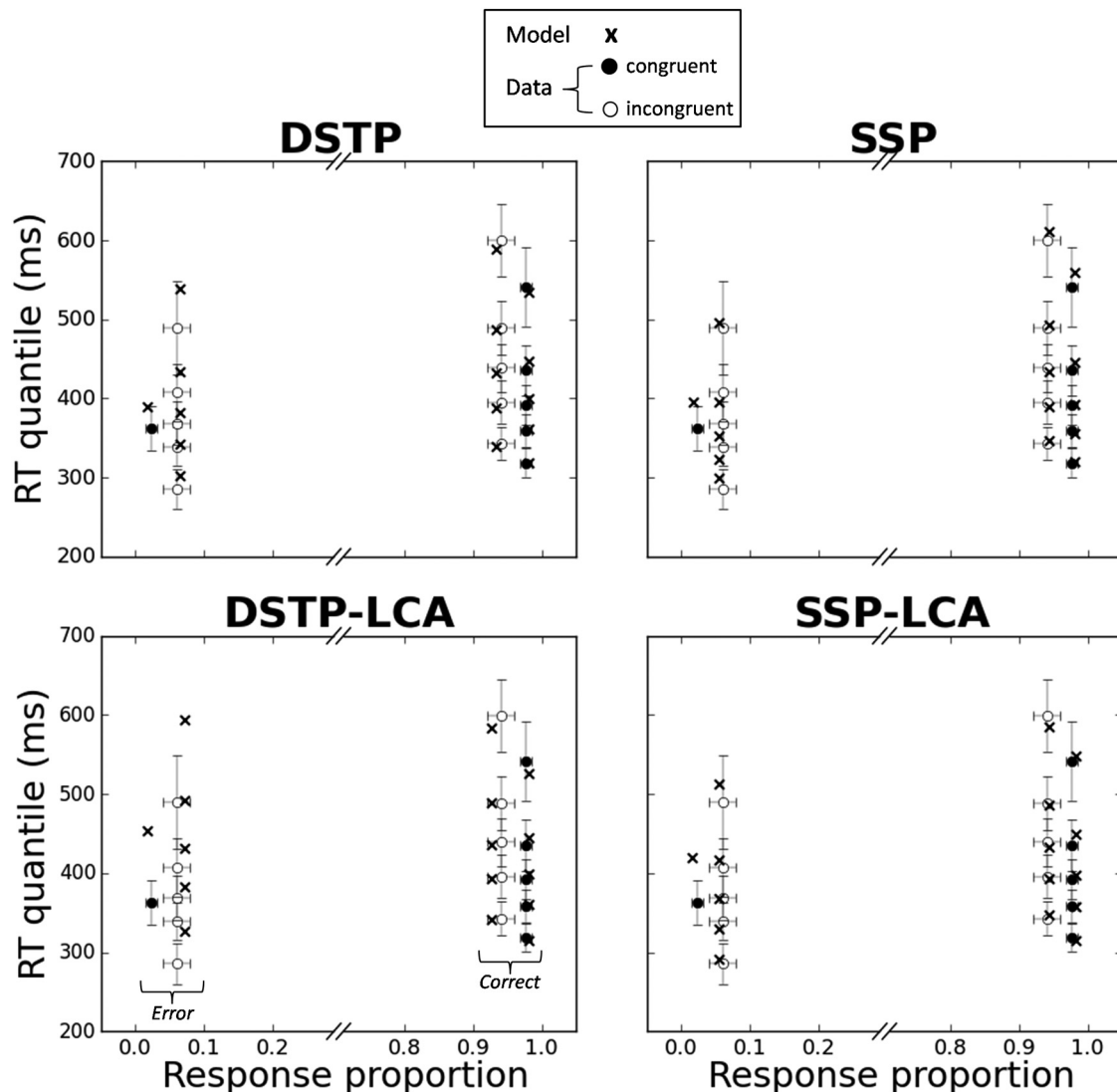
**Results**

**Behavioral data**

Responses faster than 100 ms and slower than 1500 ms were excluded from behavioral and EMG analyses (0.2% of total trials). There was a reliable congruency effect on mean RT (congruent  $M = 418$  ms; incongruent  $M = 461$  ms;  $F_{(1,11)} = 164.5$ ;  $p < 0.001$ ; Table 1). The congruency effect was also obvious on error rates (congruent  $M = 2.3\%$ ; incongruent  $M = 6\%$ ;  $F_{(1,11)} = 39.7$ ,  $p < 0.001$ ).

**Model fits**

Figure 3 shows that the DSTP, SSP, and LCA-SSP capture the main trends of the data with substantial mimicry. Most of the fitting points fall within 95% confidence intervals. The DSTP-LCA overestimates each RT quantile of errors in the incongruent condition and fails to account for the fast response capture by the



**Figure 3.** Observed (points) versus predicted (crosses) QPFs for each congruency condition averaged over the 12 participants. Each data point is accompanied by a 95% confidence interval assuming a Student’s *t* distribution and between-subject conventional SEs.

**Table 1. Error rates and mean RTs (ms) averaged across participants**

Condition	Error rate (%)	Correct RT	Error RT
Congruent trials	2.31 (1.31)	418 (48)	378 (54)
Incongruent trials	5.99 (2.99)	461 (50)	383 (53)

SDs across participants are provided in parentheses.

flankers (Gratton et al., 1988). Although the SSP-LCA has the lowest  $G^2$  statistic, the BIC slightly favors the SSP (Table 2). All models, however, overestimate the median RT for congruent trials. Inspection of individual fits reveals that the misfit of the DSTP and SSP is due to a failure of the models to capture the unusually fast median RT of errors in the congruent condition for subjects 4, 6, and 11 (Fig. 4, left). The median RT of errors in the congruent condition is generally similar to that of correct RT (Servant et al., 2014; White et al., 2011; see also the present individual data). The unusually fast median RT observed in three datasets may either be due to a high proportion of fast guesses or to the very low rate of errors that may bias the estimation of the median RT (subject 4 had the lowest rate of errors in the congruent condition 0.8%). Fast guesses could be captured by adding across-trial variability in the starting point of the accumulation

process (White et al., 2011). However, we deliberately excluded variability parameters from our analyses to focus on the primary components of the models.

Inspection of DSTP-LCA and SSP-LCA individual fits (Fig. 4, right) reveals that the models overestimate the median RT of errors in the congruent condition more strongly and frequently compared with the SSP or DSTP because the DSTP-LCA and SSP-LCA produce slower errors than correct responses in the congruent condition, a pattern rarely observed empirically. Analysis of best-fitting parameters (Table 2) reveals leak dominance (i.e., mutual inhibition much weaker than leak; for illustrations and theoretical details about this LCA processing regime, see Tsetsos et al., 2011). On average, the best-fitting mutual inhibition parameter ( $w$ ) was small for both models, showing that they approach a race (inhibition-free) model (Bogacz et al., 2006). This point will be discussed later.

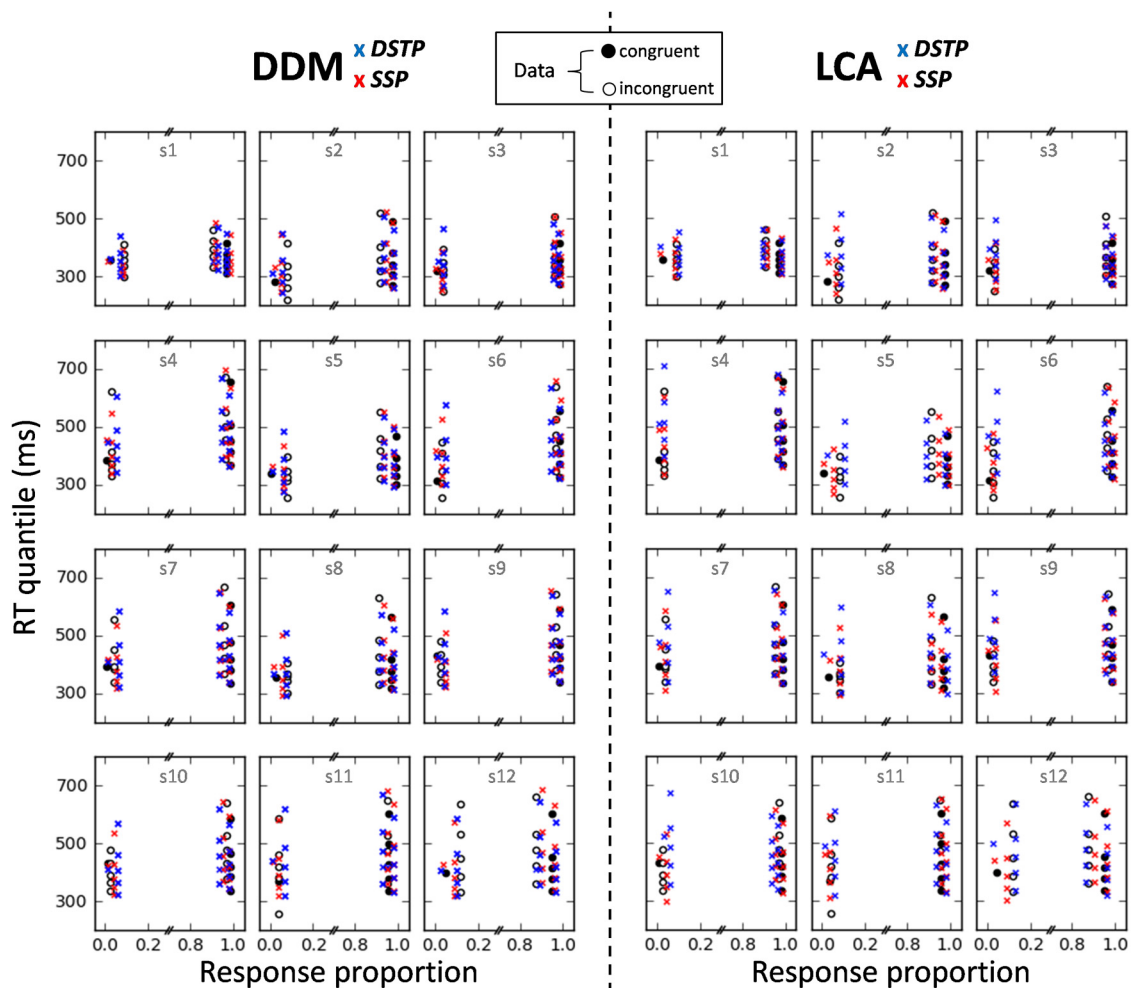
**EMG data**

The proportion of partial errors was higher in the incongruent ( $M = 26.8\%$ ) compared with the congruent ( $M = 14.4\%$ ) condition ( $F_{(1,11)} = 76.5; p < 0.001$ ; Table 3). Partial errors were slower in the incongruent ( $M = 231$  ms) compared with the

**Table 2. Best-fitting parameter values and fit statistics averaged across participants**

DSTP										
$\mu_{tar}$	$\mu_{fl}$	$a$	$c$	$Ter$	$\mu_{ss}$	$\mu_{rs2}$			$G^2$	BIC
0.128 (0.01)	0.080 (0.006)	0.149 (0.013)	0.136 (0.017)	216 (25)	0.362 (0.041)	0.915 (0.106)			79.4 (35)	130.3 (35)
SSP										
$p$	$sd_a$	$a$	$r_d$	$Ter$					$G^2$	BIC
0.356 (0.053)	2.59 (0.2)	0.117 (0.012)	0.063 (0.006)	259 (25)					83.2 (36)	119.6 (36)
DSTP-LCA										
$\mu_{tar}$	$\mu_{fl}$	$a$	$c$	$Ter$	$\mu_{ss}$	$\mu_{rs2}$	$k$	$w$	$G^2$	BIC
0.209 (0.02)	0.226 (0.047)	0.136 (0.011)	0.126 (0.009)	138 (27)	0.541 (0.041)	1.09 (0.29)	1.73 (0.97)	0.427 (0.322)	93.9 (44)	157.5 (40)
SSP-LCA										
$p$	$sd_a$	$a$	$r_d$	$Ter$			$k$	$w$	$G^2$	BIC
0.578 (0.083)	2.82 (0.412)	0.102 (0.013)	0.034 (0.005)	200 (17)			2.49 (1.38)	0.478 (0.629)	71.4 (35)	122.3 (35)

SDs across participants are shown in parentheses.  $\mu_{tar}$ , Component rate for the target;  $\mu_{fl}$ , component rate for the flankers;  $a$ , boundary separation for the response selection process;  $\mu_{ss}$ , drift rate for the target identification process;  $c$ , boundary separation for the target identification process;  $\mu_{rs2}$ , drift rate for response selection in phase 2;  $Ter$ , nondecision time;  $p$ , perceptual input;  $sd_a$ , spotlight width;  $r_d$ , spotlight shrinking rate;  $k$ , leakage;  $w$ , mutual inhibition.



**Figure 4.** Observed (points) versus predicted (crosses) QPFs for each congruency condition and participant.

congruent ( $M = 213$  ms) condition ( $F_{(1,11)} = 18.6, p < 0.005$ ), just as correction times (incongruent  $M = 149$  ms; congruent  $M = 137$  ms,  $F_{(1,11)} = 23.9, p < 0.001$ ). These findings are close to those obtained by Burle et al. (2008) with a comparable flanker task design. Those researchers reported rates of 14% and 22% for partial errors in congruent and incongruent conditions, respec-

tively. A reanalysis of their data also revealed slower partial errors and correction times for incongruent than congruent conditions ( $F_{(1,9)} = 27.5, p < 0.001$  and  $F_{(1,9)} = 13.4, p < 0.005$ ).

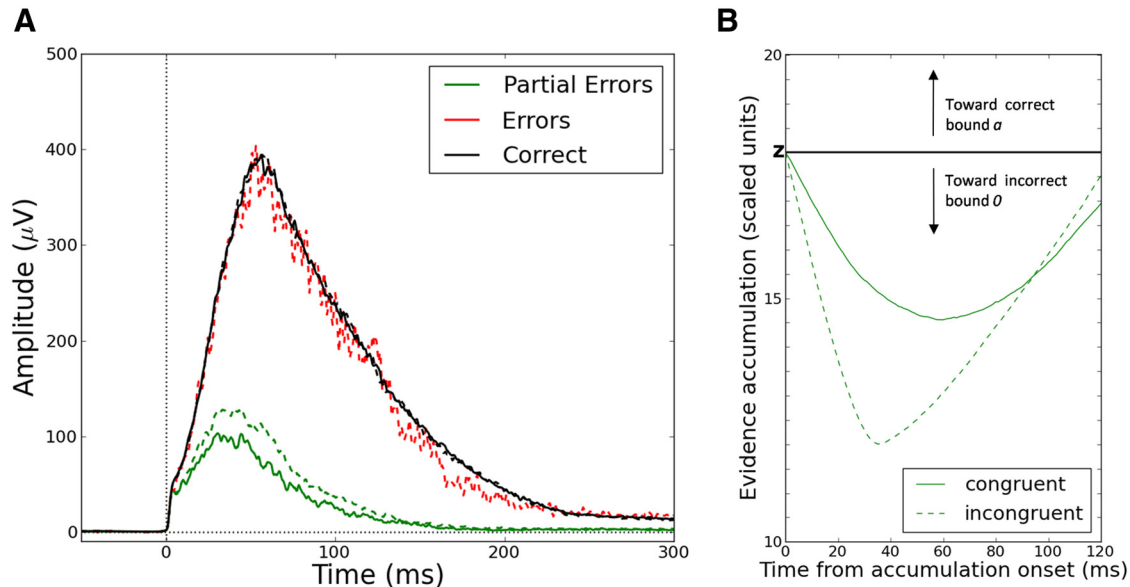
Figure 5A displays EMG bursts of partial errors and overt responses averaged over the 12 participants. EMG surfaces were submitted to an ANOVA with congruency and EMG burst type



**Table 3. EMG data and model predictions averaged across participants**

	<i>m</i>	Partial error rate (%)		Mean latency of partial errors (ms)		Mean correction time (ms)	
		Congruent	Incongruent	Congruent	Incongruent	Congruent	Incongruent
EMG		14.4 (5.9)	26.8 (10.8)	213 (35)	231 (29)	137 (34)	149 (33)
DSTP	0.0395 (0.01)	14.6 (7.0)	26.1 (9.7)	239 (31)	248 (34)	148 (23)	160 (26)
SSP	0.0287 (0.008)	11.3 (5.1)	28.0 (10.4)	271 (34)	247 (32)	143 (27)	147 (25)
DSTP-LCA	0.0851 (0.012)	10.0 (5.7)	29.3 (11.6)	305 (39)	277 (37)	64 (13)	98 (22)
SSP-LCA	0.0686 (0.011)	12.9 (5.5)	28.3 (11.2)	317 (40)	277 (38)	53 (11)	78 (12)

Partial error latencies predicted by the models incorporate an estimate of sensory encoding time. SDs across participants are shown in parentheses. *m*, Incorrect EMG bound.



**Figure 5.** *A*, Grand averages of rectified electromyographic activities time locked to EMG onset for partial errors (green lines), correct (black), and incorrect (red) overt responses. Plain lines represent congruent trials; dashed lines represent incongruent trials. Because errors are scarce in the congruent condition, the corresponding noisy data were removed for more clarity. *B*, Averaged sample paths for partial error trials (green lines) in each congruency condition time locked to the starting point of the accumulation process *z*. Sample paths were generated by the SSP diffusion model using best-fitting parameters averaged over participants and model framework shown in Figure 1*B*. Alternative models made a similar prediction (data not shown here for sake of brevity). Averaged sample paths for overt responses have a necessarily higher amplitude because they reach a decision termination bound.

(partial errors vs correct EMG bursts) as factors. Because overt errors are scarce, particularly in the congruent condition, they were removed from this analysis. The ANOVA revealed a main effect of trial type ( $F_{(1,11)} = 62.1, p < 0.001$ ). Correct EMG bursts had a much higher amplitude compared with partial errors, consistent with previous work (Rochet et al., 2014; Burle et al., 2014). The interaction between trial type and congruency was also significant ( $F_{(1,11)} = 6.7, p < 0.05$ ). Contrasts showed that the surface of partial errors was larger in the incongruent compared with the congruent condition ( $F_{(1,11)} = 9.7, p < 0.01$ ). No congruency effect was apparent on correct EMG bursts ( $F_{(1,11)} < 1$ ).

### Simulation study

A straightforward way to reconcile EMG findings of partial errors with sequential sampling models is to hypothesize two different types of bounds for the response selection process: (1) internal EMG bounds that trigger EMG activation and (2) standard bounds that terminate the decision and trigger a mechanical response. For the DSTP and SSP diffusion models, we assume EMG bounds at locations *m* (incorrect activation) and *a-m* (correct activation; Fig. 1*B*). For the DSTP-LCA and SSP-LCA models, we assume a single EMG bound *m* within each response accumulator (Fig. 1*C*). After an EMG bound is hit, evidence continues to accumulate until a decision termination bound is reached, giving the models potentiality to overcome incorrect EMG activation thanks to time-varying drift rate dynamics. Also note that there is

still residual nondesideration time after the termination bound is reached. Termination bounds cannot correspond directly to a mechanical button press because motor execution would be fully determined by the decision process. However, single-cell recordings in monkeys suggest non-decision-related motor-processing latencies (e.g., Roitman and Shadlen, 2002). Therefore, our model framework assumes that both decision-related and non-decision-related motor components contribute to EMG activity.

Diffusion models produce a partial error when a sample path hits the incorrect EMG bound *m*, but finally reaches the correct decision termination bound *a* (Fig. 1*B*). LCA models produce a partial error when the sample path of the accumulator favoring the incorrect response hits the EMG bound *m*, but that of the competing correct accumulator reaches the first decision termination bound *a* (Fig. 1*C*). Coupled with the time-varying evidence accumulation principle, EMG bounds provide a simple and intuitive explanation of EMG surface findings (Fig. 5*B*) and may account for the rate of partial errors across congruency conditions. To examine this conjecture, the models were simulated using previously obtained best-fitting parameters. For each participant, the incorrect EMG bound *m* was fit to partial error rates by minimizing the following loss function with a SIMPLEX routine:

$$Loss = \sum_{i=1}^2 (e_i - \varepsilon_i)^2$$

Where  $e$  and  $\varepsilon$  represent observed and predicted partial error rates for each congruency condition  $i$ . It should be emphasized that the parameter  $m$  was not allowed to vary between congruency conditions. A simulation study was then conducted to analyze, for each participant, predicted partial error latencies and correction times (time between first hits of the incorrect EMG bound and the correct EMG bound). Contrary to correction times, observed and predicted partial error latencies cannot be compared directly: one would need knowledge of the duration of sensory encoding time  $Te$  that affects the partial error latencies generated by the accumulator. Although  $Te$  contributes to the nondecision time parameter  $Ter$ ,  $Ter$  also incorporates a nondecision-related motor time (MT). Two methods can thus be considered to estimate  $Te$ . The first method consists in removing, for each participant  $p$  and congruency condition  $i$ , all nondecision-related motor components from  $Ter$  as follows:

$$Te_{(p,i)} = Ter_{(p)} - (\text{observed EMG } MT_{(p,i)} - \text{predicted decision-related } MT_{(p,i)}) - \text{corticomuscular delay}$$

Where the observed EMG MT is the latency from the correct EMG activation to the mechanical response (Fig. 1A) and the predicted decision-related MT is the latency between first hits of the correct EMG bound and the correct decision termination bound. Observed EMG MT and predicted decision-related MT were computed using correct responses from both pure correct and partial error trials. We also removed from  $Ter$  a necessary corticomuscular delay for the contraction of the flexor pollicis brevis. We took the upper limit of this motor conduction time estimated at 25 ms by a previous TMS/EMG study (range 20–25 ms; Burle et al., 2002). The second method consists in adjusting  $Te$  to minimize the error between observed and predicted partial error latencies for each congruency condition and participant,  $Te$  being bounded by simple physiological constraints. The lower bound for  $Te$  was defined as the minimal response latency of neurons from the primary visual cortex ( $\sim 40$  ms; Poghosyan and Ioannides, 2007). The upper bound was defined as  $Ter$  minus the corticomuscular delay. Following the generally accepted hypothesis that flanker congruency does not affect sensory encoding processes (Kornblum et al., 1990; Ridderinkhof et al., 1995; Liu et al., 2008; Hübner et al., 2010; White et al., 2011),  $Te$  was also constrained to be fixed across congruency conditions. Because both methods yielded similar conclusions, we only report results from the first method where  $Te$  is a prediction of the models, not a fit to data.

### Simulation results

Table 3 shows best-fitting values for the incorrect EMG bound  $m$  along with predicted partial error rates/latencies and correction times averaged over participants. All models approximately captured the observed partial error rates across congruency conditions, with the best fit provided by the DSTP. However, predicted partial error latencies and correction times across congruency conditions differ as a function of processing assumptions. We will first present simulation results from the DDM and then the LCA models.

### DSTP and SSP simulation results

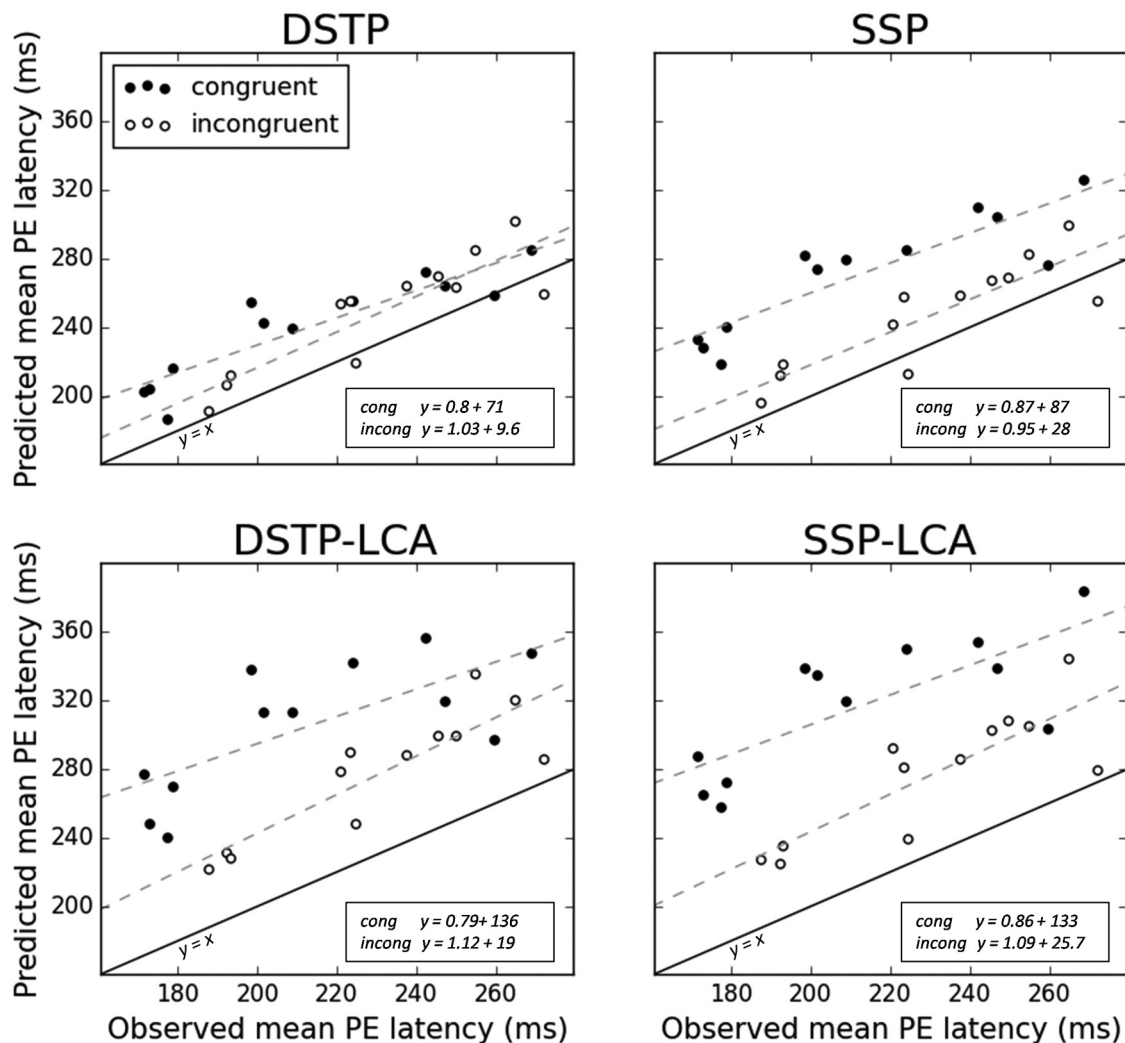
Figure 6 illustrates the relationship between observed and predicted mean partial error latencies (incorporating our estimate of  $Te$ ) for each congruency condition and participant. The intersub-

ject correlation between these two variables was very high both for congruent (DSTP:  $r = 0.9$ ,  $p < 0.001$ ; SSP:  $r = 0.88$ ,  $p < 0.001$ ) and incongruent (DSTP:  $r = 0.88$ ,  $p < 0.001$ ; SSP:  $r = 0.86$ ,  $p < 0.001$ ) conditions. Importantly, this correlation was systematically higher compared with that obtained when removing  $Te$  (DSTP: congruent  $r = 0.57$ ,  $p = 0.06$ , incongruent  $r = 0.41$ ,  $p > 0.1$ ; SSP: congruent  $r = 0.5$ ,  $p = 0.1$ , incongruent  $r = 0.4$ ,  $p > 0.1$ ; data not shown), which reassured us that our strategy for estimating  $Te$  is pertinent. The regression slopes were close to one (Fig. 6, insets), particularly for the incongruent condition in which partial error rates are higher and latencies better estimated. The mismatch between model predictions and experimental data is mainly due to a fixed offset. Mean partial error latencies are overestimated by 26 ms (congruent condition) and 17 ms (incongruent) on average for the DSTP and by 58 ms (congruent) and 16 ms (incongruent) on average for the SSP (Table 3). The more important mismatch of the SSP in the congruent condition is due to a failure of the model to capture the observed congruency effect on partial error latencies, as detailed below.

The DSTP predicted faster partial errors in the congruent compared with the incongruent condition ( $t_{(11)} = 5.6$ ,  $p < 0.001$ , paired sample  $t$  test; Fig. 7, Table 3), in agreement with observed data, whereas the SSP predicted the reverse pattern ( $t_{(11)} = -12.1$ ,  $p < 0.001$ ). These opposite predicted effects are not explained by differences in  $Te$ ; on average, we found approximately similar  $Te$  estimates across congruency conditions for each model (DSTP: congruent  $M = 154$  ms, incongruent  $M = 153$  ms; SSP: congruent  $M = 192$  ms, incongruent  $M = 194$  ms), consistent with the hypothesis that flanker congruency does not affect sensory encoding processes (Kornblum et al., 1990; Ridderinkhof et al., 1995; Liu et al., 2008; Hübner et al., 2010; White et al., 2011). The opposite model predictions are instead explained by differences in drift rate dynamics. Averaged model fits show that flanker interference is originally larger in the SSP ( $v(t) = -0.244$ ) compared with the DSTP ( $\mu_{\text{tar}} - \mu_{\text{fl}} = 0.047$ ). Moreover, the drift rate for the response selection process of the DSTP starts off strong in congruent trials ( $\mu_{\text{tar}} + \mu_{\text{fl}} = 0.208$ ) and gets 4.4 times stronger after stimulus identification ( $\mu_{\text{rs}2} = 0.915$ ). Contrary to the SSP, a partial error must occur very early, because the second phase of response selection makes it unlikely that the diffusion would dip below the incorrect EMG bound  $m$ .

Both DDM models predicted faster correction times for congruent than incongruent conditions (DSTP:  $t_{(11)} = 6.2$ ,  $p < 0.001$ ; SSP:  $t_{(11)} = 3.9$ ,  $p < 0.005$ ), in agreement with observed EMG data (Table 3). Figure 8 displays observed and predicted cumulative distribution functions of correction times averaged over participants for each congruency condition. Cumulative distribution functions were constructed by computing correction time quantiles (0.1, 0.3, 0.5, 0.7, 0.9). Both models captured fairly well the distribution of correction times across congruency conditions. It should be emphasized that the gray crosses of Figure 8 are not a fit to data, which makes the performance of diffusion models even more impressive.

Predicted decision-related MTs are smaller than observed EMG MTs (Fig. 9, compare the ordinate of data points across subplots), leaving some residual latency for nondecision-related motor components, consistent with our model framework. If predicted decision-related MTs have physiological validity, then they should correlate with observed EMG MT. We thus computed the intersubject correlation between predicted decision-related MTs and observed EMG MTs for correct responses. This correlation proved to be very high (SSP: congruent  $r = 0.87$ ,  $p < 0.001$ , incongruent  $r = 0.86$ ,  $p < 0.001$ ; DSTP: congruent  $r = 0.9$ ,

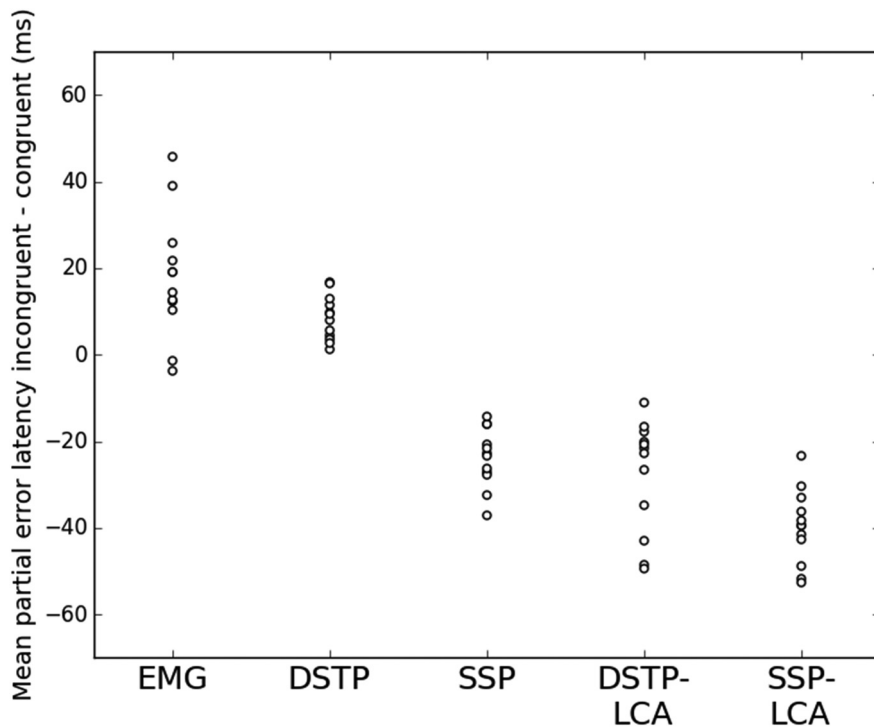


**Figure 6.** Predicted versus observed mean partial error latency (both in ms) for each participant and congruency condition. Model predictions are not a fit to data and incorporate an estimate of sensory encoding time (see text for details). Also shown are lines of best fit for each congruency condition (dashed lines, equation provided in the insets) and the ideal  $y = x$  line (plain line).

$p < 0.001$ , incongruent  $r = 0.9$ ,  $p < 0.001$ ). In addition to this relationship, the model framework makes another nontrivial prediction: when the incorrect EMG bound  $m$  increases (i.e., moves further away from the termination bounds), the rate of partial errors increases as well as the decision-related MT for correct responses. In other words, for each congruency condition, a positive correlation between the rate of partial errors and the EMG MT of correct responses is predicted. Figure 9 shows that this association holds in our data, and a high positive correlation was found both for congruent ( $r = 0.80$ ,  $p < 0.005$ ) and incongruent ( $r = 0.79$ ,  $p < 0.005$ ) conditions. Similar results were obtained by correlating partial error rates to EMG MT of correct responses from pure correct trials (congruent:  $r = 0.80$ ,  $p < 0.005$ ; incongruent:  $r = 0.81$ ,  $p < 0.005$ ) and to EMG MT of correct responses from partial error trials (congruent:  $r = 0.80$ ,  $p < 0.005$ ; incongruent:  $r = 0.79$ ,  $p < 0.005$ ) separately.

Although our extended diffusion model framework explains several aspects of behavioral and EMG flanker task data, it does not specify how EMG activity stops. We assumed that EMG activity starts when the weight of evidence for one response alternative exceeds a criterion (EMG bound). One could hypothesize that EMG activity stops when the weight of evidence steps back and returns below this criterion. This assumption appears inap-

propriate. In our simulations, we observed that sample paths often meander round EMG bounds (resulting in several hits). Taken literally, such meandering diffusion would trigger a volley of EMG bursts. However, such trials are scarce; only a small proportion of observed trials contain multiple EMG bursts (classified as “other” trials in the present study). We ran a final simulation of the models using best-fitting parameters averaged over participants to determine the percentage of partial error trials in which the sample path meander round the correct EMG bound  $a-m$  after the partial error. Those percentages were abnormally high (DSTP: congruent = 60%, incongruent = 58%; SSP: congruent = 78%; incongruent = 78%). This finding may suggest that the magnitude of diffusion noise  $s$  is too high. This parameter serves as a scaling parameter and has traditionally been fixed to 0.1 (Ratcliff and Smith, 2004), but this value does not rely on any neurophysiological principle. A lower amount of noise would make the diffusion process more deterministic, meaning that the sample path would not hover around the EMG bound. Alternatively, EMG activity may stop when the sample path reaches the opposite EMG bound. This hypothesis predicts that the end of the partial error matches the onset of the corrective EMG burst. However, results from Burle et al. (2008) suggested the existence of a gap between the two events. In the current data, we detected the end



**Figure 7.** Magnitude (in ms) of observed and predicted congruency effect on partial error latencies for each participant. Model predictions incorporate an estimate of sensory encoding time.

of each observed partial error by visual inspection and analyzed the cumulative distribution function of the delta difference between the corrective EMG burst onset latency and the end of the partial error. Figure 10 shows that delta is always positive, even for the lowest quantile, refuting the matching hypothesis.

**DSTP-LCA and SSP-LCA simulation results**

The intersubject correlation between observed and predicted mean partial error latencies (incorporating our estimate of  $T_e$ ) was high for each LCA model and congruency condition (DSTP-LCA: congruent  $r = 0.71, p < 0.01$ , incongruent  $r = 0.87, p < 0.001$ ; SSP-LCA: congruent  $r = 0.76, p < 0.005$ , incongruent  $r = 0.83, p < 0.001$ ), although the correlation coefficient in the congruent condition was smaller compared with diffusion models (Fig. 6). Correlations were systematically higher than those obtained when removing  $T_e$  (DSTP-LCA: congruent  $r = 0.38, p > 0.1$ , incongruent  $r = 0.29, p > 1$ ; SSP-LCA: congruent  $r = 0.53, p = 0.08$ , incongruent  $r = 0.59, p < 0.05$ ).

Both models strongly overestimated partial error latencies in the congruent condition (DSTP-LCA: magnitude of overestimation congruent  $M = 92$  ms, incongruent  $M = 46$  ms; SSP-LCA: congruent  $M = 104$  ms, incongruent  $M = 46$  ms) and thus predicted faster partial errors in the incongruent compared with the congruent condition ( $t_{(11)} = -7.4, p < 0.001$  and  $t_{(11)} = -13.5, p < 0.001$  for the DSTP-LCA and SSP-LCA respectively; Fig. 7, Table 3), opposite to empirical findings. Estimated  $T_e$ s were approximately similar across congruency conditions for the SSP-LCA (congruent:  $M = 138$  ms; incongruent:  $M = 139$  ms), but not for the DSTP-LCA (congruent:  $M = 84$  ms; incongruent  $M = 69$  ms). In addition to the congruency effect on  $T_e$ , notice that the incorrect response accumulator of the DSTP-LCA also generates faster partial errors in the incongruent ( $M = 208$  ms) compared with the congruent condition ( $M = 221$  ms). Although this result appears at odds with the DSTP diffusion model, the architecture

and drift rate dynamics of the two models are different. In particular, when a correct target is identified in the congruent condition, the drift rate of the DSTP-LCA’s incorrect response accumulator drops at 0 (and not  $-\mu_{rs2}$ ). The DSTP-LCA can thus generate slow partial errors in the congruent condition, contrary to the diffusion model version.

Analysis of LCA model dynamics revealed another important problem. In a substantial number of partial error trials, the sample path of the accumulator favoring the incorrect response hit the EMG bound, whereas the sample path of the correct accumulator had already engaged correct EMG activity and was located between  $m$  and  $a$ . This phenomenon was more frequent in congruent (DSTP-LCA: 42%; SSP-LCA: 42%) than in incongruent (DSTP-LCA: 21%; SSP-LCA: 22%) trials. Trials with overlapping correct and incorrect EMG bursts are rare in the observed data (Fig. 10), particularly when the incorrect EMG burst starts after the correct one. In this latter case (classified as “other” trials), the correction time would be negative, which is meaningless. We thus excluded such trials from our correction time analysis.

Even in this favorable case, the DSTP-LCA and SSP-LCA strongly underestimated correction times (Fig. 8, Table 3). These failures of the LCA models suggest that the coactivation of the two competing accumulators is problematic. Response coactivation cannot occur in the DDM framework because evidence for one response alternative is evidence against the other alternative. The only way in the LCA to reduce coactivation is to increase the mutual inhibition parameter  $w$ . This contrasts with the best fits to the behavioral data, which required a low amount of mutual inhibition (Table 2). For thoroughness, we simulated the LCA models with high mutual inhibition ( $w = 6$ ) to determine whether they could capture the behavioral and EMG data. At the EMG level, the models predicted slower correction times and a reduced proportion of abnormal partial error trials. At the behavioral level, however, they overestimated accuracy and the skew of RT distributions for correct responses. Therefore, it appears that the LCA models cannot simultaneously account for the behavioral and EMG data without drastically changing the level of mutual inhibition. In contrast, the DDMs, particularly DSTP, accounted for both behavioral and EMG data with the same assumption of perfect inhibition.

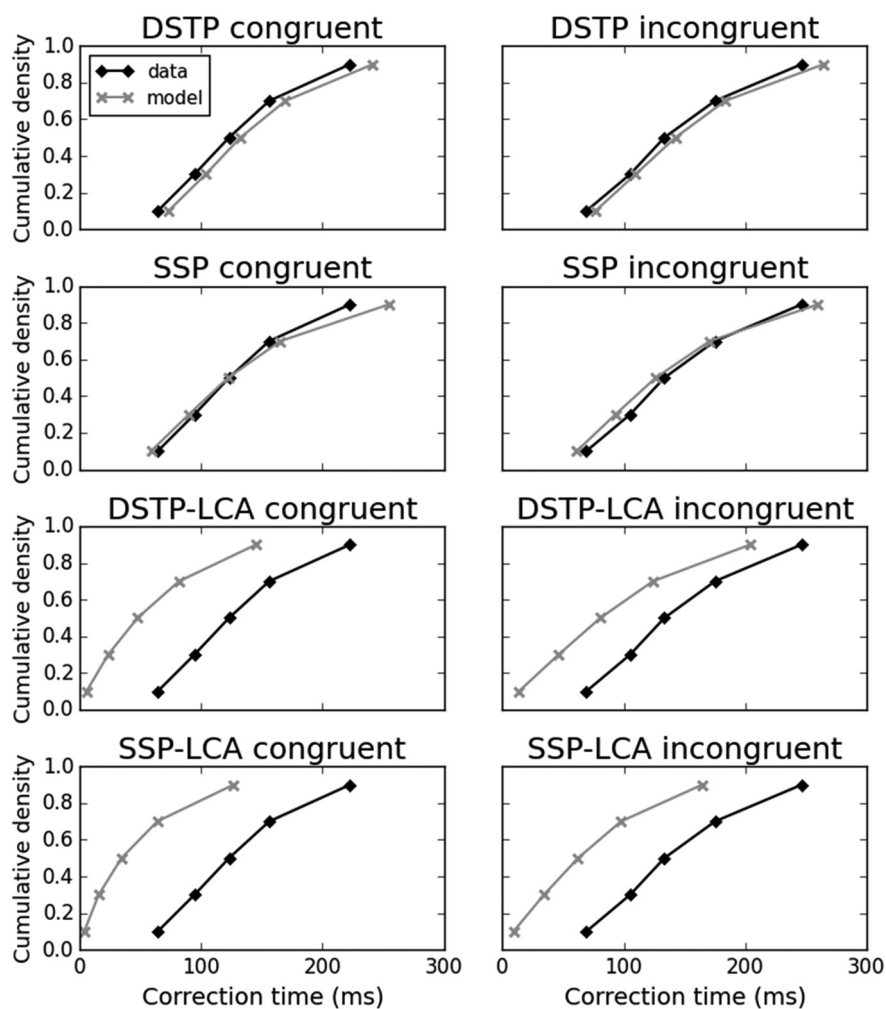
Beyond these important failures, we found some similarities between LCA and DDM model performance. Predicted decision-related MTs were smaller than observed EMG MTs (Fig. 9), and the two variables were strongly correlated (DSTP-LCA: congruent  $r = 0.91, p < 0.001$ , incongruent  $r = 0.82, p < 0.005$ ; SSP-LCA: congruent  $r = 0.91, p < 0.001$ , incongruent  $r = 0.9, p < 0.001$ ). The LCA models also predict a positive correlation between the rate of partial errors and the EMG MT of correct responses for each congruency condition (Fig. 9). Finally, we found that sample paths of the correct response accumulator often hover around the EMG bound after a partial error (DSTP-LCA: congruent = 40%, incongruent = 48%; SSP-LCA: congruent = 55%; incongruent = 67% of partial error trials). Because both the

LCA and DDM models were simulated with  $s = 0.1$ , this result further supports the hypothesis that the magnitude of noise is too high.

## Discussion

Most decisions that we make build upon multiple streams of sensory evidence and control mechanisms are needed to filter out irrelevant information. The DDM has recently been extended to account for decisions about relevance (Hübner et al., 2010; White et al., 2011). Although this new generation of models incorporates attention filters that weight sensory evidence in a flexible and goal-directed way, it preserves the longstanding hypothesis that motor execution occurs only once a decision termination bound is reached. At face value, the mere presence of EMG partial errors during the decision process (Coles et al., 1985; Burle et al., 2014) seems to challenge this view. Consistent with this, we found a flanker congruency effect on partial error EMG surfaces, suggesting that part of motor activity is underpinned by the decision process. We thus assumed that the brain engages motor activity when the weight of evidence for one response alternative exceeds a first criterion (EMG bound). Importantly, evidence continues to accumulate until a standard decision termination bound is reached, offering the potentiality to overcome incorrect EMG activation thanks to time-varying drift rate dynamics. Under this assumption, a DDM implementing a discrete improvement of attentional selectivity (DSTP) captured the rate, EMG surface, and latency of partial errors across congruency conditions along with the speed of the correction process. The model also predicted a positive correlation between partial error rates and EMG MTs, which was empirically verified, supporting the hypothesis that part of EMG activity reflects up-to-date information about the decision.

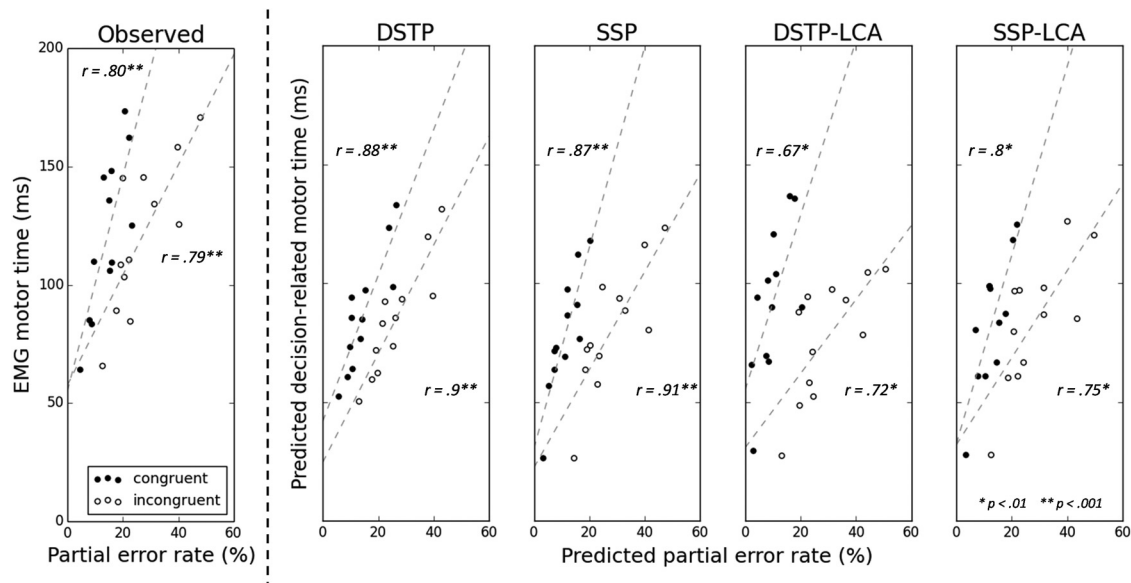
Within the theoretical framework of a standard DDM, Kiani and Shadlen (2009) have recently demonstrated that a confidence signal is encoded by the same neurons that represent the decision variable in the parietal cortex. Confidence is defined as the log odds probability favoring one alternative over the other and depends upon accumulated evidence (see Fig. 4B in Kiani and Shadlen, 2009). Its main properties are that log odds correct decreases when decision time increases and accumulated evidence for the incorrect alternative increases. We may speculate that the brain engages motor activity in an adaptive way when the accumulated evidence reaches a sufficiently high degree of confidence. In case of partial errors, evidence accumulation would rapidly drift toward the incorrect decision bound, resulting in high confidence for the incorrect response. If the weight of evidence starts to favor the correct response (after attentional selection), then the confidence signal would drop dramatically, leaving space for correction. Our constant EMG bounds may be a simplified realization of this idea. However, it is unclear how



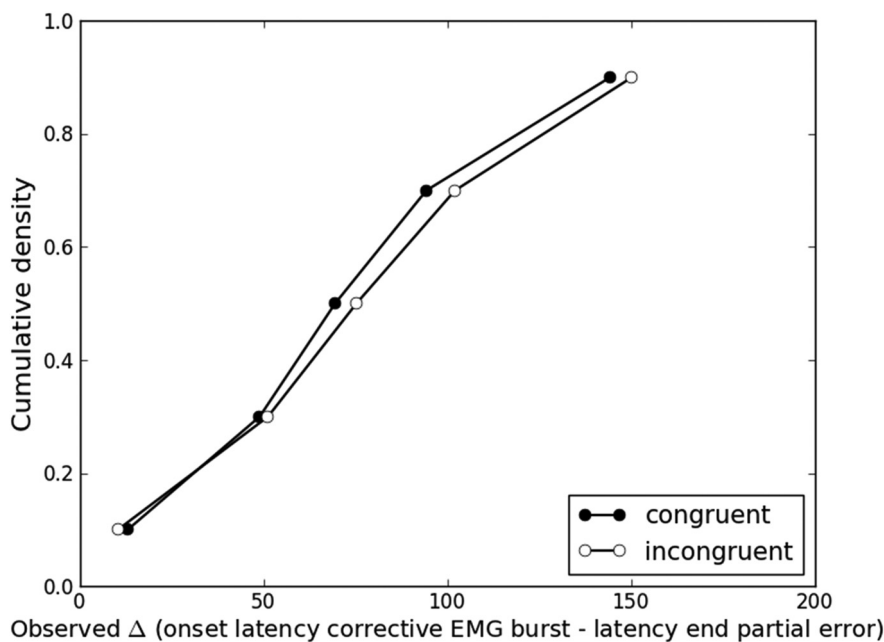
**Figure 8.** Cumulative distribution functions of observed (diamonds) and predicted (crosses) correction times (in ms) for each congruency condition averaged over the 12 participants. Predicted correction times are not a fit to data.

EMG bounds incorporated into a single-model accumulator map onto activities of numerous accumulator neurons that encode confidence signals in the parietal cortex (Schall, 2004). The *e pluribus unum* model developed by Zandbelt et al. (2014) sheds light on this issue. The model demonstrates that ensembles of numerous and redundant accumulator neurons can generate a RT distribution similar to that predicted by a single accumulator, provided that accumulation rates are moderately correlated and RT is not determined by the lowest and highest accumulation rates. Decision-bound invariance across trials also arises from the accumulation neural network dynamics under similar conditions. Therefore, the use of a single-model accumulator coupled with our assumption of EMG and decision-bound invariance are pertinent to model the present behavioral and EMG data if these broad and reasonable constraints also hold with time-varying accumulation rates.

An alternative formalism is the change-of-mind diffusion model proposed by Resulaj et al. (2009). According to this view, partial errors would reflect an initial decision that is subsequently revised. The model assumes that evidence accumulation carries on in the postdecision period. The initial decision is revised if a “change-of-mind bound” is reached within a time period or reaffirmed otherwise. The model builds upon a time-constant diffusion process and was not designed to account for decisions



**Figure 9.** Left, Observed EMG MT (in ms) for correct trials versus observed partial error rate for each congruency condition and participant. Also shown are lines of best fit (dashed lines) and Pearson's  $r$  coefficient correlations. Right, Predicted decision-related MT for correct trials (in ms) versus predicted partial error rate from each model, congruency condition, and participant.



**Figure 10.** Cumulative distribution functions of the delta latency difference (in ms) between the corrective EMG burst and the end of the partial error for each congruency condition averaged over the 12 participants.

based on multiple streams of sensory evidence. In addition, it was meant to explain overt changes of mind in decision making revealed by hand trajectories and did not make specific hypotheses about covert EMG activity. Apart from the change-of-mind extension, the framework preserves standard diffusion model assumptions and implicitly assumes that motor execution occurs once a decision termination bound is reached. Therefore, it is unclear whether it can explain the congruency effect on partial error EMG surfaces and the strong positive correlation between partial error rates and EMG MTs. In addition, it has been shown recently that partial errors are consciously detected in only 30% of cases (Rochet et al., 2014). It remains to be determined whether a commitment to a choice can occur without conscious

access (see Shadlen and Kiani, 2013, for a discussion about the links between evidence accumulation and consciousness). Further work is required to determine whether a change-of-mind model provides a better explanation of EMG empirical findings than our EMG-bound model.

By linking EMG activity to sensory evidence accumulation, our study introduces a new methodology to test latent assumptions of formal decision-making models. Partial errors provide an intermediate measure of evidence accumulation at the single trial level and appear useful to constrain model dynamics. We put into test two extended DDMs (DSTP and SSP) that strongly mimic each other on behavioral grounds despite qualitatively different processing assumptions (Hübner et al., 2010; White et al., 2011). These models made opposite predictions with regard to the direction of the flanker congruency effect on partial error latencies. Further analyses suggested that the diffusion noise, a scaling parameter traditionally fixed at 0.1 (Ratcliff and Smith, 2004), may be too high. We compared these DDMs with another sequential sampling model class (LCA) extended with similar attentional mechanisms. Interestingly, one of these extended LCA models (SSP-LCA) captured the behavioral data equally well with a very low amount of mutual inhibition between competing accumulators, approaching a race (inhibition-free) model (Bogacz et al., 2006). However, the model generated abnormal EMG patterns and strongly underestimated the latency of the partial error correction process, suggesting that a high degree of mutual inhibition between competing accumulators is necessary to explain the EMG data, in contradiction with the optimal fitting of behavioral data. Finally, we observed that all tested models systematically

overestimate partial error latencies across participants, with the smallest overestimation ( $\sim 20$  ms on average) provided by the DSTP. This slight overestimation is unlikely to be due to an underestimation of the corticomuscular delay, given that we took the upper known limit of the motor conduction time (Burle et al., 2002). Rather, it may be due to another residual processing stage not related to sensory encoding or motor execution (Gold and Shadlen, 2007) or to a problem in decision-making dynamics. These results encourage the use of EMG to better constrain sequential sampling models.

An important finding arising from our approach is that part of EMG activity is determined by the decision process. Beyond sequential sampling models, this finding challenges traditional continuous flow models of information processing. For example, McClelland's (1979) cascade model framework assumes response execution to be "a discrete event that adds the duration of a single discrete stage to the time between the presentation of the stimulus and the registration of the overt response" (p. 291). Contrary to this view, our results suggest a continuity between decision making and response engagement that necessarily implies a representation of the decision variable in the cortical motor areas (Cisek, 2006). Electrophysiological studies in monkeys have revealed that decision formation and motor preparation share similar neural networks (Gold and Shadlen, 2007). For example, when monkeys have to decide the net direction of random dot motion and communicate their decision by means of a saccade, neurons in the frontal eye field represent the accumulated motion evidence (Gold and Shadlen, 2000). Similar findings have been reported in humans (Heekeren et al., 2008; Donner et al., 2009; O'Connell et al., 2012). In particular, a convergence of techniques (electroencephalography, transcranial magnetic stimulation) in between-hand two-choice RT demonstrate that the developing activation of the contralateral primary motor cortex (M1) involved in the required response is accompanied by a developing inhibition over the ipsilateral M1 (Burle et al., 2004). Whether these activities reflect the accumulated weight of sensory evidence for each response alternative requires more investigation.

This work also represents an advance in the horizontal integration of processes involved in simple decision making through the use of cognitive modeling. The standard DDM focuses only on the decision process and does not account for where the evidence comes from (encoding and task-specific processing) nor how the response was executed. The SSP and DSTP models advanced upon this framework by specifying how perceptual encoding and attentional control drive the time-varying decision evidence that is fed into the DDM framework. Finally, the present work extends these models to specify how the evolving decision evidence relates to activation of the motor system to produce the response. The final result is a model that links together different processes involved in translating the stimulus into a response and successfully accounts for behavioral and neurophysiological data from an array of stimulus and decision manipulations (Hübner et al., 2010; White et al., 2011).

## Notes

Supplemental material for this article is available at [http://sites.univ-provence.fr/lnc/supplementary\\_material\\_Servant\\_et\\_al.pdf](http://sites.univ-provence.fr/lnc/supplementary_material_Servant_et_al.pdf). This material has not been peer reviewed.

## References

Bogacz R, Brown E, Moehlis J, Holmes P, Cohen JD (2006) The physics of optimal decision making: a formal analysis of models of performance in two-alternative forced-choice tasks. *Psychol Rev* 113:700–765. [CrossRef Medline](#)

- Broadbent DE (1958) Perception and communication. London: Pergamon.
- Burle B, Bonnet M, Vidal F, Possamaï CA, Hasbroucq T (2002) A transcranial magnetic stimulation study of information processing in the motor cortex: relationship between the silent period and the reaction time delay. *Psychophysiology* 39:207–217. [CrossRef Medline](#)
- Burle B, Vidal F, Tandonnet C, Hasbroucq T (2004) Physiological evidence for response inhibition in choice reaction time tasks. *Brain Cogn* 56:153–164. [CrossRef Medline](#)
- Burle B, Roger C, Allain S, Vidal F, Hasbroucq T (2008) Error negativity does not reflect conflict: a reappraisal of conflict monitoring and anterior cingulate cortex activity. *J Cogn Neurosci* 20:1637–1655. [CrossRef Medline](#)
- Burle B, Spieser L, Servant M, Hasbroucq T (2014) Distributional reaction time properties in the Eriksen task: marked differences or hidden similarities with the Simon task? *Psychon Bull Rev* 21:1003–1010. [CrossRef Medline](#)
- Cisek P (2006) Integrated neural processes for defining potential actions and deciding between them: a computational model. *J Neurosci* 26:9761–9770. [CrossRef Medline](#)
- Coles MG, Gratton G, Bashore TR, Eriksen CW, Donchin E (1985) A psychophysiological investigation of the continuous flow model of human information processing. *J Exp Psychol Hum Percept Perform* 11:529–553. [CrossRef Medline](#)
- Deutsch JA, Deutsch D (1963) Attention: some theoretical considerations. *Psychol Rev* 70:80–90. [CrossRef Medline](#)
- Donner TH, Siegel M, Fries P, Engel AK (2009) Buildup of choice-predictive activity in human motor cortex during perceptual decision making. *Curr Biol* 19:1581–1585. [CrossRef Medline](#)
- Eriksen BA, Eriksen CW (1974) Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & Psychophysics* 16:143–149. [CrossRef](#)
- Eriksen CW, St James JD (1986) Visual attention within and around the field of focal attention: a zoom lens model. *Percept Psychophys* 40:225–240. [CrossRef Medline](#)
- Forstmann BU, Wagenmakers EJ, Eichele T, Brown S, Serences JT (2011) Reciprocal relations between cognitive neuroscience and formal cognitive models: opposites attract? *Trends Cogn Sci* 15:272–279. [CrossRef Medline](#)
- Gold JI, Shadlen MN (2000) Representation of a perceptual decision in developing oculomotor commands. *Nature* 404:390–394. [CrossRef Medline](#)
- Gold JI, Shadlen MN (2007) The neural basis of decision making. *Annu Rev Neurosci* 30:535–574. [CrossRef Medline](#)
- Gratton G, Coles MG, Sirevaag EJ, Eriksen CW, Donchin E (1988) Pre- and poststimulus activation of response channels: a psychophysiological analysis. *J Exp Psychol Hum Percept Perform* 14:331–344. [CrossRef Medline](#)
- Heekeren HR, Marrett S, Ungerleider LG (2008) The neural systems that mediate human perceptual decision making. *Nat Rev Neurosci* 9:467–479. [CrossRef Medline](#)
- Heitz RP, Schall JD (2012) Neural mechanisms of speed-accuracy tradeoff. *Neuron* 76:616–628. [CrossRef Medline](#)
- Hübner R, Töbel L (2012) Does attentional selectivity in the Flanker Task improve discretely or gradually? *Front Psychol* 3:434. [CrossRef Medline](#)
- Hübner R, Steinhauser M, Lehle C (2010) A dual-stage two-phase model of selective attention. *Psychol Rev* 117:759–784. [CrossRef Medline](#)
- Kiani R, Shadlen MN (2009) Representation of confidence associated with a decision by neurons in the parietal cortex. *Science* 324:759–764. [CrossRef Medline](#)
- Kornblum S, Hasbroucq T, Osman A (1990) Dimensional overlap: cognitive basis for stimulus-response compatibility—a model and taxonomy. *Psychol Rev* 97:253–270. [CrossRef Medline](#)
- Liu YS, Holmes P, Cohen JD (2008) A neural network model of the Eriksen task: reduction, analysis, and data fitting. *Neural Comput* 20:345–373. [CrossRef Medline](#)
- McClelland J (1979) On the time relations between of mental processes: an examination of systems of processes in cascade. *Psychol Rev* 86:287–330. [CrossRef](#)
- Meyer DE, Osman AM, Irwin DE, Yantis S (1988) Modern mental chronometry. *Biol Psychol* 26:3–67. [CrossRef Medline](#)
- Nelder JA, Mead R (1965) A simplex method for function minimization. *Comput J* 7:308–313. [CrossRef](#)
- O'Connell RG, Dockree PM, Kelly SP (2012) A supramodal accumulation-

- to-bound signal that determines perceptual decisions in humans. *Nat Neurosci* 15:1729–1735. [CrossRef Medline](#)
- Peirce JW (2007) PsychoPy—Psychophysics software in Python. *J Neurosci Methods* 162:8–13. [CrossRef Medline](#)
- Poghosyan V, Ioannides AA (2007) Precise mapping of early visual responses in space and time. *Neuroimage* 35:759–770. [CrossRef Medline](#)
- Purcell BA, Heitz RP, Cohen JY, Schall JD, Logan GD, Palmeri TJ (2010) Neurally constrained modeling of perceptual decision making. *Psychol Rev* 117:1113–1143. [CrossRef Medline](#)
- Purcell BA, Schall JD, Logan GD, Palmeri TJ (2012) From salience to saccades: multiple-alternative gated stochastic accumulator model of visual search. *J Neurosci* 32:3433–3446. [CrossRef Medline](#)
- Ratcliff R (1978) A theory of memory retrieval. *Psychol Rev* 85:59–108. [CrossRef](#)
- Ratcliff R, McKoon G (2008) The diffusion decision model: theory and data for two-choice decision tasks. *Neural Comput* 20:873–922. [CrossRef Medline](#)
- Ratcliff R, Smith PL (2004) A comparison of sequential sampling models for two-choice reaction time. *Psychol Rev* 111:333–367. [CrossRef Medline](#)
- Ratcliff R, Hasegawa YT, Hasegawa RP, Smith PL, Segraves MA (2007) Dual diffusion model for single-cell recording data from the superior colliculus in a brightness-discrimination task. *J Neurophysiol* 97:1756–1774. [Medline](#)
- Ratcliff R, Philiastides MG, Sajda P (2009) Quality of evidence for perceptual decision making is indexed by trial-to-trial variability of the EEG. *Proc Natl Acad Sci U S A* 106:6539–6544. [CrossRef Medline](#)
- Resulaj A, Kiani R, Wolpert DM, Shadlen MN (2009) Changes of mind in decision-making. *Nature* 461:263–266. [CrossRef Medline](#)
- Ridderinkhof KR, van der Molen MW, Bashore TR (1995) Limits on the application of additive factors logic: violations of stage robustness suggest a dual-process architecture to explain flanker effects on target processing. *Acta Psychologica* 90:29–48. [CrossRef](#)
- Rochet N, Spieser L, Casini L, Hasbroucq T, Burle B (2014) Detecting and correcting partial errors: Evidence for efficient control without conscious access. *Cogn Affect Behav Neurosci* 14:970–982. [CrossRef Medline](#)
- Roitman JD, Shadlen MN (2002) Response of neurons in the lateral intraparietal area during a combined visual discrimination reaction time task. *J Neurosci* 22:9475–9489. [Medline](#)
- Sadja P, Philiastides MG, Heekeren H, Ratcliff R (2011) Linking neuronal variability to perceptual decision making via neuroimaging. In: *The dynamic brain: an exploration of neuronal variability and its functional significance* (Ding M, Glanzman DL, eds), pp 214–232. New York: OUP.
- Schall JD (2004) On building a bridge between brain and behavior. *Annu Rev Psychol* 55:23–50. [CrossRef Medline](#)
- Schurger A, Sitt JD, Dehaene S (2012) An accumulator model for spontaneous neural activity prior to self-initiated movement. *Proc Natl Acad Sci U S A* 109:E2904–E2913. [CrossRef Medline](#)
- Schwarz G (1978) Estimating the dimension of a model. *Ann Stat* 6:461–464. [CrossRef](#)
- Servant M, Montagnini A, Burle B (2014) Conflict tasks and the diffusion framework: Insight in model constraints based on psychological laws. *Cogn Psychol* 72:162–195. [CrossRef Medline](#)
- Shadlen MN, Kiani R (2013) Decision making as a window on cognition. *Neuron* 80:791–806. [CrossRef Medline](#)
- Stauder G, Flachenecker C, Daumer M, Wolf W (2001) Onset detection in surface electromyographic signals: a systematic comparison of methods. *EURASIP J Adv Sig Pr* 2:67–81.
- Storn R, Price K (1997) Differential evolution: a simple and efficient heuristic for global optimization over continuous spaces. *Journal of Global Optimization* 11:341–359. [CrossRef](#)
- Tsetsos K, Usher M, McClelland JL (2011) Testing multi-alternative decision models with non-stationary evidence. *Front Neurosci* 5:63. [Medline](#)
- Usher M, McClelland JL (2001) The time course of perceptual choice: the leaky, competing accumulator model. *Psychol Rev* 108:550–592. [CrossRef Medline](#)
- Van Boxtel GJ, Geraats LH, Van den Berg-Lenssen MM, Brunia CH (1993) Detection of EMG onset in ERP research. *Psychophysiology* 30:405–412. [CrossRef Medline](#)
- van Maanen L, Brown SD, Eichele T, Wagenmakers EJ, Ho T, Serences J, Forstmann BU (2011) Neural correlates of trial-to-trial fluctuations in response caution. *J Neurosci* 31:17488–17495. [CrossRef Medline](#)
- White CN, Brown S, Ratcliff R (2012) A test of Bayesian observer models of processing in the Eriksen flanker task. *J Exp Psychol Hum Percept Perform* 38:489–497. [Medline](#)
- White CN, Ratcliff R, Starns JJ (2011) Diffusion models of the flanker task: discrete versus gradual attentional selection. *Cogn Psychol* 63:210–238. [CrossRef Medline](#)
- White CN, Mumford JA, Poldrack RA (2012) Perceptual criteria in the human brain. *J Neurosci* 32:16716–16724. [CrossRef Medline](#)
- Winer BJ (1971) *Statistical principles in experimental design: design and analysis of factorial experiments*. New York: McGraw-Hill.
- Zandbelt B, Purcell BA, Palmeri TJ, Logan GD, Schall JD (2014) Response times from ensembles of accumulators. *Proc Natl Acad Sci U S A* 111:2848–2853. [CrossRef Medline](#)